EAI Endorsed Transactions
on Internet of Things
Research Article EAI.EU

Weed detection with Improved Yolov 7

Mingkang Peng[1], Wuping Zhang[1,*], Fuzhong Li[1], Qiyuan Xue[1], Jialiang Yuan [1], Peipu An[1]

[1]College of Software, Shanxi Agricultural University, Taigu, Jinzhong , Shanxi, China

Abstract

INTRODUCTION: An improved Yolov7 model.
OBJECTIVES: To solve the weed detection and identification in complex field background.
METHODS: The dataset was enhanced by online data enhancement, in which the feature extraction, feature fusion and feature point judgment of weed image were carried out by Yolov7 to predict the weed situation corresponding to the prior box. In the enhanced feature extraction part of Yolov7, CBAM, an attention mechanism combining channel and space, is introduced to improve the attention of the algorithm to weeds and strengthen the characteristics of weeds.
RESULTS: The mean average precision (mAP) of the improved algorithm reached 91.15%, which was 2.06% higher than that of the original Yolov7 algorithm. Compared with the current mainstream target detection algorithms Yolox, Yolov5l, Fster RCNN, Yolov4-tiny and Yolov3, the mAP value of the improved algorithm increased by 4.35, 4.51, 5.41, 19.77 and 20.65 percentage points. Weed species can be accurately identified when multiple weeds are adjacent.
CONCLUSION: This paper provides a detection model based on Yolov7 for weed detection in the field, which has a good detection effect on weed detection, and lays a research foundation for intelligent weeding robot and spraying robot.

Keywords: weed identification, deep learning, attention mechanism, yolov7

Received on 20 June 2023, accepted on 20 July 2023, published on 31 July 2023

doi: 10.4108/eetiot.v9i3.3468

1. Introduction

Weed damage is the main factor affecting agricultural development[1]. In farmland, weeds compete with crops for water, nutrients and spatial location, which in turn affects the yield and quality of crops[2-3]. How to reasonably control weeds has become a major problem in agricultural development. At present, artificial pesticide spraying devices and drone spraying are the main ways to control weeds, which will have a bad impact on human body and environment, and even endanger human life[4-8]. Weed detection is the basis of intelligent weed control and intelligent equipment research[9-11], and it is also the difficulty[12]. Therefore, it is necessary to study the accurate detection of weeds.

Previous weed detection methods mostly use traditional machine learning methods to extract image features manually, and only use simple phenotypic features such as color and traits, which are easily affected by factors such as human and environmental background, resulting in low recognition accuracy and poor robustness[13-16].

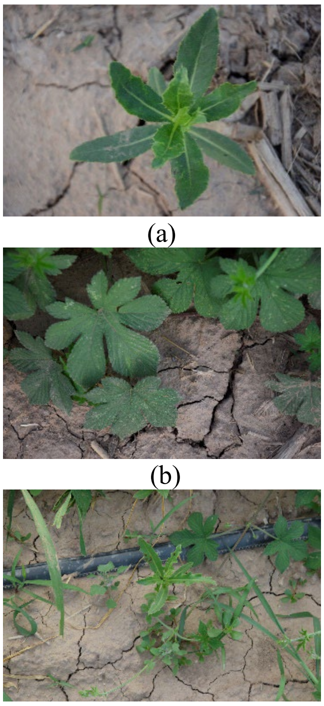
2. Methodology

2.1 Data Preparation

The data required for the experiment were collected from the test site of the west station of Shanxi Agricultural University and the test site of the pastoral station of Shanxi Agricultural University in Taigu District, Jinzhong City, Shanxi Province. The collected objects were the associated weeds information of soybeans, corn and other crops, including nine kinds of weed images of thistle grass, humulus scandens, daylily, summer solstice grass,

*Corresponding author. Email: zwping@126.com

amaranth, quinoa, phoenix, solanum and glass grass. The acquisition equipment is Canon SLR camera EOS70D and Nikon SLR camera D3300. The maximum image resolution is 6000 * 4000. The collected images are characterized by individual weeds, multiple same weeds and multiple different weeds. Some samples are shown in Figure 1.



(a) (b) (c)

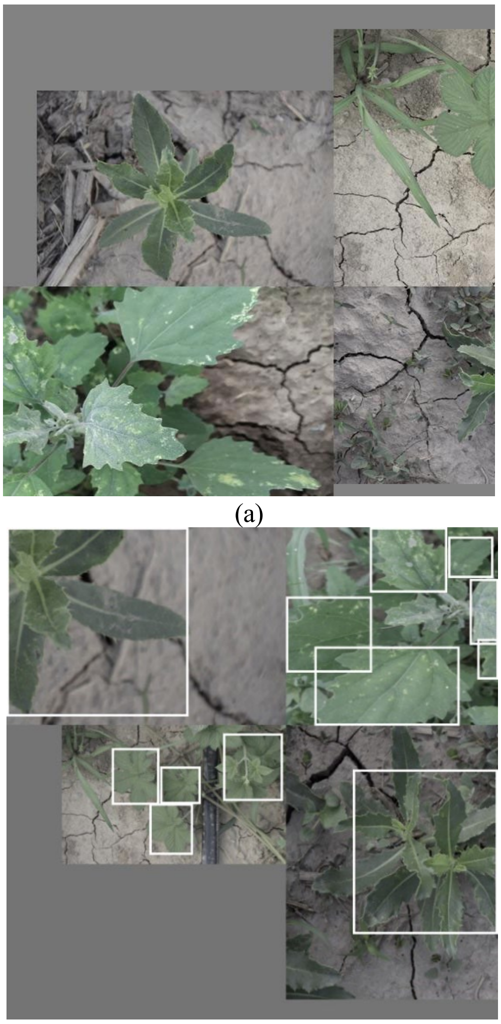
Figure 1. Examples of sample. (a)Single weed, (b)Multiple identical weeds,(c)Multiple weeds.

2.2 Data Pre-Processing

After screening to remove duplicate and blurred images, the data set used 400 thistles, 350 humulus scandens, 230 bowl flowers, 160 summer solstice grass, 700 amaranth, 700 quinoa, 270 lanceolate, 150 solanum, 210 borage, a total of 3170. According to the training set: validation set : test set for the ratio of 8 : 1 : 1 to establish a data set, in order to ensure reliable data, using a random method to divide the data set, the data sets do not coincide with each other. The training set is used to fit the model, the classification model is trained by the set hyperparameters, the validation set is used to adjust the model parameters, and the test set is used to evaluate the performance of the model. We use the annotation tool LabelIMG to annotate images and generate Pascal VOC dataset format.

2.3 Data Augmentation

Data enhancement is divided into online data enhancement and offline data enhancement. In this paper, we use online data enhancement method. Mosaic ' s idea is to cut four images randomly and then stitch them into one image as training data. During the training process, each step has a 50 % probability to use Mosaic data enhancement, and a 50 % probability to mix-up data enhancement of Mosaic enhanced images. At the same time, the training set is randomly flipped and the color gamut is transformed. Gray bars are added to the redundant parts of the image, so that each epoch training picture is different. This method can greatly enrich the data set and enhance the robustness and generalization of the model. Compared with offline data enhancement, online data enhancement does not need to save the enhanced data, which saves a lot of storage space and has strong flexibility. The image effect after data enhancement is shown in Figure 2.



(a) (b)

Figure 2. Effects of image data augmentation. (a) Picture splicing cutting, (b)Automatically add labels.

2.4 Experimental Environment

The training parameters of the training process used in the experiment are shown in Table 1.

Table 1. Experimental platform configuration

Configuration	Configuration
Operating System	Windows10
CPU	Intel(R)Core(TM)I7-5930K
Random Access Memory	48G
GPU	NVIDIA TITAN X(Pascal)
Python	3.8
Pytorch	1.12.1

2.5 Train Parameters

The training set is used to train the model. The fp16 mixed precision training is used to train 200 epochs, and the first 50 epochs are frozen for training. The learning rate of the model is set to 0.01, and the SGD optimizer is used. The momentum parameter in the optimizer is set to 0.937, the weight attenuation coefficient is set to 0.0005, the Batch Size is set to 16, and the cosine annealing attenuation strategy is used to adjust the learning rate.

2.6 Evaluation indicators

In order to represent the performance of the model, it was decided to use Average Precison, Mean of Precision (mAP), F1-score, parameters, and model memory usage as model evaluation metrics. AP is related to Precision (P) and recall (R), which is the area under the curve of PR curve. AP value can better reflect the effect of the model. The specific calculation formula is:

$$P = \frac{TP}{TP + FP} \times 100\% \tag{1}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{2}$$

$$AP = \int_0^1 P(R)dR \tag{3}$$

$$mAP = \frac{\sum AP}{N} \tag{4}$$

$$F1 = 2 \times \frac{PR}{P + R} \tag{5}$$

In the formula, TP is the number of positive samples correctly predicted; FP is the number of samples whose error prediction is positive; FN is the number of samples whose false prediction is negative; N is the number of predicted samples.

2.7 Related Network

In this section, the details of the introduction of Yolov7 into the attention mechanism are described.

2.7.1 Improved Yolov7 framework

Yolov7 is a one stage algorithm designed by the original Yolov4 team. As the latest algorithm of Yolo series, Yolov7 has the advantages of higher accuracy and faster speed than Yolov5 series under the same training amount. Yolov7 exceeds the currently known detector between 5FPS and 160FPS[17-19]. The improved Yolov7 algorithm structure is shown in Figure 3.

Yolov7 structure can be divided into three parts: BackBone, FPN, Yolo Head. After the weed data is entered, the feature layer is first obtained in the BackBone section through convolution, batch normalization, and activation function processing. In order to identify weeds more accurately, Yolov7 is improved. The improvement is mainly to add CBAM after each MCB in the backbone feature network to enhance the attention to weed features, so as to improve the accuracy of weed identification.

In the BackBone section, get the feature set for the input image data. The feature set is input into the FPN part for feature fusion. In the experiment, the data set contains complex images, in which the size and characteristics of the target object are different. By using the shallow features in FPN, a large feature map can be obtained, so as to identify simple targets. For example, pointed leaf grass belongs to Gramineae, Humulus scandens belongs to broadleaf grass, and the leaf morphology of the two is very different. Using deep features can obtain small target features and identify complex targets. In the process of image processing, FPN continuously down-sampled the feature points of the target, and the image size became smaller and smaller. After up-sampling, the image size became larger and larger, so as to obtain more useful information. Through feature upsampling and downsampling, feature fusion is finally achieved to obtain three enhanced feature layers.

The enhanced feature layer of the weed obtained by the BackBone and FPN parts is the feature point set, and the feature points are judged by Yolo Head to determine whether the prior frame on the feature point corresponds to the weed.

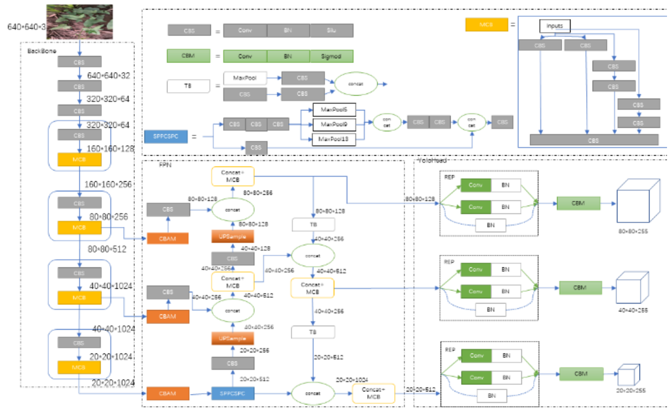


Figure 3. Improved Yolov7 Algorithm structure diagram

2.7.2 CBAM attention mechanism

Because the weed shooting environment is too complex, the accuracy of target detection is low. To reduce the influence of noise, attention mechanism CBAM was introduced to improve the attention to the key features of weeds. The attention mechanism originated in the field of machine translation, mainly for image recognition tasks. In 2018, a lightweight attention mechanism CBAM was proposed. CBAM is a mixed attention mechanism composed of channels and space.

The channel attention module enters the image F and passes through the global maximum pooling and the global average pooling to obtain two vectors. The Relu activation function is used to fuse the two vectors through the multi-layer perceptron and use the addition method for data fusion. Finally, the activation function sigmoid activation operation obtains the channel attention vector Mc (Channel Attention) process as shown in Figure 4.

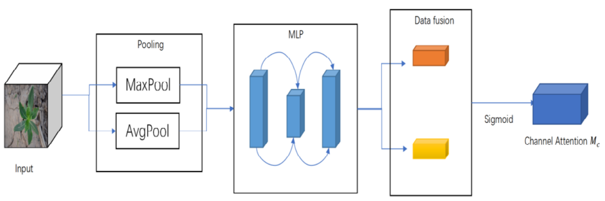


Figure 4. Channel attention process

The calculation formula of channel attention vector M_c is:

$$M_c(F) = \sigma\Big(MLP\big(AvgPool(F)\big) + MLP\big(MaxPool(F)\big)\Big)$$
$$= \sigma\Big(W_1\left(W_0\left(F_{avg}^c\right)\right) + W_1\left(W_0\left(F_{max}^c\right)\right)\Big) \qquad (6)$$

σ is sigmoid function, W_0 and W_1 are weight size of MLP, AvgPool is average pooling, MaxPool is max pooling, F_{avg}^c is average pooling feature, F_{max}^c is max pooling feature.

Spatial module obtains the feature map along the channel dimension after the global maximum pooling and global average pooling to obtain different feature description operator, convolution kernel is 7 * 7 convolution Sigmoid activation function, finally get the spatial attention vector (Spatial branch), the process is shown in Figure 5.

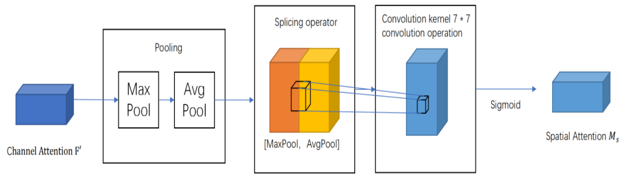


Figure 5. Spatial attention process

The calculation formula of spatial attention vector M_s is:

$$M_s(F) = \sigma(f^{7\times7}([AvgPool(F); MaxPool(F)]))$$
$$= \sigma\left(f^{7\times7}\left(\left[F_{avg}^s; F_{max}^s\right]\right)\right) \qquad (7)$$

where $f^{7\times7}$ is a convolution kernel of size 7 x 7.

3 Experimental Results Analysis

3.1 Experimental results

Using Yolov7-CBAM to train the training set according to the adjusted parameters, we get the loss function Figure 6(a) and mAP Figure6(b). The loss function is used to describe the difference between the predicted value and the true value of the model. The smaller the loss function value, the smaller the difference between the predicted value and the true value. The loss function of Yolov7-CBAM includes training loss and verification loss. It can be seen from the graph that both training loss and verification loss converge to a lower value, which proves that the improved Yolov7 has good convergence ability and high robustness of the model. From Figure 6(b), it can be concluded that in the training of the model, the mAP value of the first 50 epoch models is improved rapidly, and the model tends to be stable after training to 100 epochs. When training to 200 epochs, the mAP value of Yolov7-CBAM is stable at about 0.91.

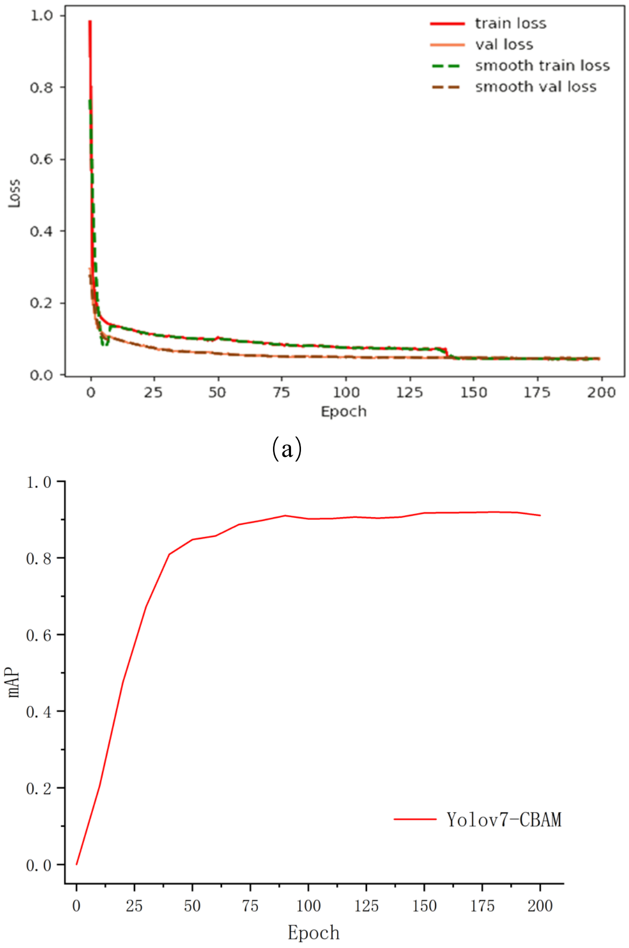


Figure 6. Train curve, (a)Loss function curve, (b)mAP curve.

3.2 Comparison of detection results using different attention mechanisms

In order to verify the effectiveness of adding attention mechanism, different attention mechanisms are compared with no attention mechanism. The comparison results are shown in Table 2. The average accuracy of Yolov7 category without attention mechanism is 89.09 %, the average accuracy of Yolov7 category with CA attention mechanism is 90.62 %, the average accuracy of Yolov7 category with CBAM attention mechanism is 91.15 %, and the average accuracy of Yolov7 category with SE attention mechanism is 88.79 %. After using CBAM mechanism, the problem of difficult identification of Solanum nigrum is solved without affecting the recognition accuracy of other categories, and the mAP value is increased by 2 percentage points.

Table 2. Average classification accuracy of different attention mechanisms %

Types	Yolov7	Yolov7-CA	Yolov7-CBAM	Yolov7-SE
Crisium	89.05	97.35	97.87	97.54
Calystegia	88.27	95.45	95.26	95.99
Humulus	89.48	88.10	85.24	80.95
Pteris ensiformis	90.16	89.23	93.18	90.46
Chenopodium album	80.74	84.78	85.47	82.14
Cynoglossum furcatum	94.29	90.71	91.86	93.25
Solanum nigrum	85.93	85.15	94.55	83.24
Amaranth	88.31	86.19	83.85	85.13
Lagopsis supina	95.56	92.68	93.08	90.21
Average	89.09	90.62	91.15	88.79

3.3 Comparison of detection results using different model

Table 3 and Figure7 shows the comparison results of different model. It can be seen from Table 3 that the mAP value and F1 value of Yolov7-CBAM algorithm in training weed data are higher than those of Yolov7, Yolox, Faster RCNN, Yolov5l, Yolov4-tiny and Yolov3. Although Yolov7 takes up more memory and parameters, Yolov7-CBAM algorithm is selected because the mAP value and F1 value can better reflect the accuracy of the model.

Table 3. Comparison results of performance indicators of different Models

Models	mAP/%	F1-score/%	Model Memory/MB	Parameters
Yolov7	89.09	79.55	143	37.23M
Yolov7-CBAM	91.15	80.11	143	37.56M
Yolox	86.80	77.55	34.3	8.94M
Faster RCNN	86.64	68.22	108	28.35M
Yolov5l	85.74	70.44	176	46.18M
Yolov4-tiny	71.38	60.00	22.5	5.89M
Yolov3	70.50	44.33	235	61.56M

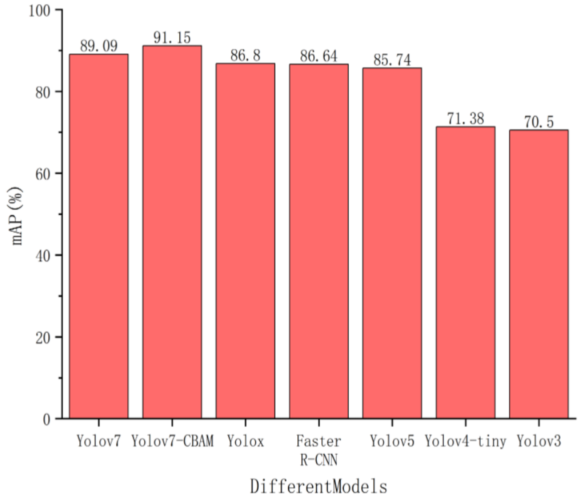
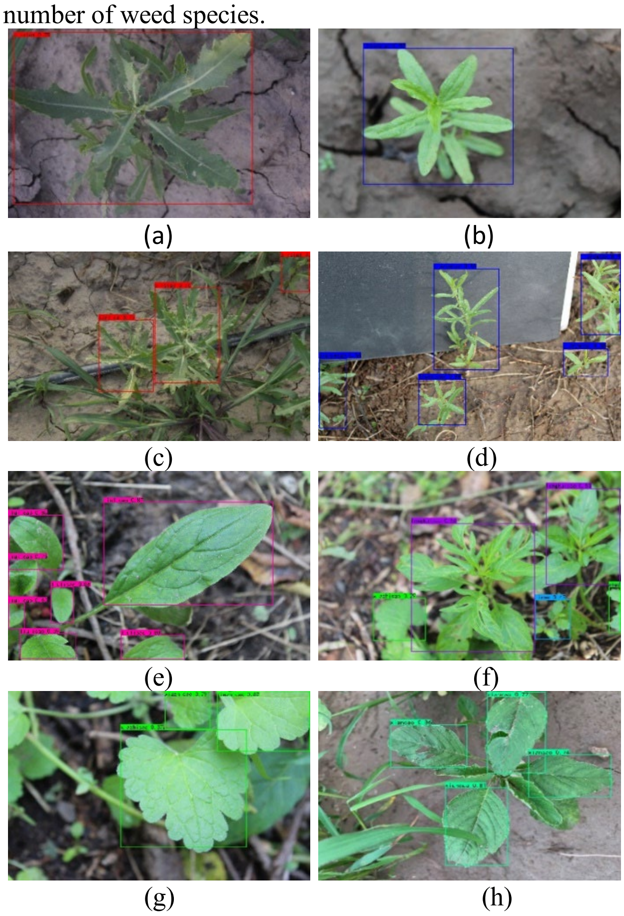


Figure 7. Accuracy of models

3.4 Yolov7-CBAM detecting result

The trained model was used to detect weeds, and three different background environments of single plant weeds, single plant weeds and multiple plants weeds were selected for testing, as shown in Figure8. When multiple weeds are adjacent, the optimized Yolov7 algorithm can still accurately identify each weed, which is basically not affected by the complex environmental background and the number of weed species.



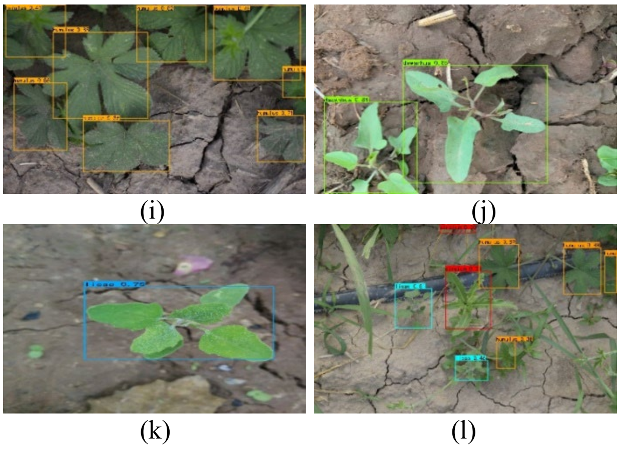


Figure 8. Result of detection. (a)Cirsium, (b)Pteris ensiformis, (c)Multiple Cirsium, (d)Multiple Pteris ensiformis, (e)Cynoglossum furcatum, (f)Multiple weeds, (g)Lagopsis supina, (h)Amaranth, (i)Multiple humulus, (j)Calystegia, (k)Chenopodium album, (l)Multiple weeds.

4. Discussion

4.1 Development of deep learning of agriculture

Deep learning has become a popular method for image classification in recent years due to its high detection accuracy and strong robustness[20]. Classical deep learning algorithms include R-CNN[21] and Mask R-CNN[22].This paper uses one stage target detection algorithm. One stage target detection algorithm has the advantage of fast recognition speed[23], including SSD[24], RetinaNet[25], Yolov3[26] and so on.Jiang[27] et al. proposed a graph convolutional network(GCN) method based on CNN features. The GCN-Resnet-101 method achieved 97.8 %, 99.37 %, 98.93 % and 96.51 % recognition accuracy on four different data sets. Chen[28] et al. used Yolov4 algorithm with attention mechanism and feature fusion to detect sesame associated weeds, and the detection accuracy reached 96.16 %. Jin[29] et al.applied Yolov3 to weeding in vegetable fields, identified vegetable crops, and classified other objects as weeds. The detection accuracy of Yolov3 for vegetables reached more than 97 %, avoiding the problem of low detection accuracy due to the wide variety of weeds. Tang[30] et al. constructed a weed recognition model based on K-means combined with CNN. K-means pre-training was used to replace the random initialization of CNN weights. The K-means pre-training method achieved 92.89 % accuracy for weed data. Zhao Hui[31] et al.introduced the attention mechanism on the DenseNet network to enhance the adaptability of different weed types through regularization, and the average recognition accuracy of corn seedlings and six associated weeds reached 98.63 %.

4.2 Limitations

The Yolov7-CBAM algorithm proposed in this paper has a significant effect in weed recognition. Compared with other algorithms, it greatly improves the detection accuracy. It can accurately identify each weed in a complex environment, but there are still some limitations. First of all, the image data set used in this paper is self-collected, and the variety is relatively single. Due to various environmental factors, the weed characteristics of each area may be different, so in the future research, we can collect more weed data sets and expand the data set. Secondly, the Yolov7l model occupies a large space, and future research needs to further improve our model and reduce the space occupied by the model.

5. Conclusion

Combining Yolov7 algorithm and attention mechanism, a Yolov7-CBAM algorithm more suitable for weed detection is proposed. The average accuracy of weeds is 91.15 %, which is better than Yolov7 with other attention mechanisms, YOlov7 without improvement, other Yolo algorithms and Faster R-CNN. It is proved that the introduction of CBAM attention mechanism into Yolov7 helps to reduce the influence of environmental noise and strengthen the key features of weeds, thereby improving the accuracy of weed recognition.

Through online data enhancement, the generalization and robustness of the algorithm are enhanced, and the accurate detection of weeds is realized in the scene of single weed, single multiple weeds and multiple weeds. It has practical value for the development of weed robots and spraying robots.

References

[1] Zhao Yuxin, Yang Huiming. Effects of crop pattern, tillage practice and water and fertilizer management on weeds and their control mechanisms. Acta Prataculturae Sinica. 2015; 24(8):199-210.

[2] Pan Junfeng, Wan Kaiyuan, Tao Yong, et al. Ecological weed control strategy based on farmland nutrient management. Plant Protection. 2014; 40(3):5-9+36.

[3] Cui cui, Tang yin. Effect of sowing rate of wheat on the weed community and wheat yield. Journal of Southwest University (Natural Science Edition). 2011; 33(12):12-17.

[4] Madhukar S.Chavan. Automatic Arial Vehicle Based Pesticides Spraying System for Crops. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 2019; 8(11).

[5] Giles D., Billing R. Deployment and Performance of a Uav for Crop Spraying. Chemical Engineering Transactions (CET Journal), 2015; 44.

[6] Daniel E. Martin, Wayne E. Woldt, Mohamed A. Latheef. Effect of Application Height and Ground Speed on Spray Pattern and Droplet Spectra from Remotely Piloted Aerial Application Systems. Drones. 2019; 3(4).

[7] Baharuddin Mohd Rafee B, Sahid Ismail B, Noor Mohamad Azhar B Mohd, et al. Pesticide risk assessment: A study on inhalation and dermal exposure to 2,4-D and paraquat among Malaysian paddy farmers. Journal of environmental science and health. Part. B, Pesticides, food contaminants, and agricultural wastes. 2011; 46(7).

[8] Christos A. Damalas, Spyridon D. Koutroubas. Farmers' Exposure to Pesticides: Toxicity Types and Ways of Prevention. Toxics. 2016; 4(1).

[9] Hu Lian, Liu Hailong, He jie, et al. Research progress and prospect of intelligent weeding robot. Journal of South China Agricultural University. 2023; (1):1-15.

[10] Fu Leiyang, Li shaowen, Zhang le, et al. Research progress on Field Weeding Robots:A Review. ROBOT. 2021; 43(6):751-768.

[11] Li Yong, Guo Zhiqiang, Shuang Feng, et al. Key technologies of machine vision for weeding robots: A review and benchmark. Computers and Electronics in Agriculture. 2022; 196.

[12] Yuan Hongbo, Zhao Nudong, Cheng Man. Review of Weeds Recognition Based on Image Processing. Transactions of the Chinese society for Agricultural Machinery. 2020; 51(S2):323-334.

[13] Yan Qing, Liang Dong, Zhang Dongyan. Recognition of weed in corn field based on supervised locally linear embedding algorithm. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE). 2013; 29(14):171-177.

[14] Li Hui, Qi Lijun, Zhang Jianhua. Recognition of Weed during Cotton Emergence Baesd on Principal Component Analysis and Support Vector Machine. Transactions of the Chinese society for Agricultural Machinery. 2012; 43(09):184-189+196.

[15] Wu Lanlan, Liu Jiangying, Wen Youxian. Weed Identification Method Based on SVM in the Corn Field. Transactions of the Chinese society for Agricultural Machinery. 2009; 40(01):162-166.

[16] Cheng Beibei, ERIC T. MATSON. A Feature-Based Machine Learning Agent for Automatic Rice and Weed Discrimination. Artificial intelligence and soft computing: 14th International Conference, ICAISC 2015, Zakopane, Poland, June 14-18. 2015; Proceedings, p.I.:Springer, 2015:517-527.

[17] Wang,C.Y. Bochkovskiy, A.Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv 2022, arXiv:2207.02696.

[18] Bochkovskiy, A.; Wang,C.Y.;Liao,H.Y.M.Yolov4:Optimal speed and accuracy of object detection.arXiv 2020, arXiv:2004.10934

[19] Jiang Kailin, Xie Tianyu, Yan Rui, et al. An Attention Mechanism-Improved YOLOv7 Object Detection Algorithm for Hemp Duck Count Estimation. Agriculture. 2022; 12(10).

[20] Ji Shouling, Du Tianyu, Deng Shuiguang, et al. Robustness Certification Research on Deep Learning Models: A Survey. Chinese Journal of Computers. 2022; 45(01):190-206.

[21] Ren Shaoqing, He Kaiming, Girshick Ross, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.IEEE transactions on pattern analysis and machine intelligence. 2017; 39(6).

[22] He Kaiming, Gkioxari Georgia, Dollar Piotr, et al. Mask R-CNN.[J].IEEE transactions on pattern analysis and machine intelligence. 2020; 42(2).

[23] Li Kecen, Wang Xiaoqiang, Lin Hao, et al. Survey of One-Stage small object detection methods in deep learning. Journal of Frontiers of Computer Science and Technology. 2022; 16(01):41-58.

[24] Wei Liu, Dragomir Anguelov, Dumitru Erhan, et al. SSD: Single Shot MultiBox Detector. CoRR,2015,abs/1512.02325.
[25] Tsung-Yi Lin, Piotr Dollár, Ross B. Girshick, et al. Feature Pyramid Networks for Object Detection. CoRR,2016,abs/1612.03144.
[26] Redmon, J., Farhadi, A. Yolov3: An incremental improvement. 2018; arXiv preprint arXiv:1804.02767.
[27] Jiang Honghua, Zhang Chuanyin, Qiao Yongliang, et al. CNN feature based graph convolutional network for weed and crop recognition in smart farming. Computers and Electronics in Agriculture. 2020; 174(C).
[28] Chen Jiqing, Wang Huabin, Zhang Hongdu, et al. Weed detection in sesame fields using a YOLO model with an enhanced attention mechanism and feature fusion. Computers and Electronics in Agriculture. 2022; 202.
[29] Jin Xiaojun, Sun Yanxia, Che Jun, et al. A Novel Deep Learning-Based Method for Detection of Weeds in Vegetables. Pest management science. 2022; 78(5).
[30] Tang JingLei, Wang Dong, Zhang ZhiGuang, et al. Weed identification based on K-means feature learning combined with convolutional neural network. Computers and Electronics in Agriculture. 2017; 135.
[31] Zhao Hui, Cao Yuhang, Yue Yongjin, et al. Field weed recognition based on improved DenseNet. Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE). 2021; 37(18):136-142.